

# K-Nearest Neighbors (K-NN) Classification

for EIE417 course by Suneesh Jacob A.

## Question

You are given a dataset with two input features  $x_1$  and  $x_2$ , and a continuous output  $y$ . The dataset is as follows:

S. No.	$x_1$	$x_2$	$y$
1	4	1	0
2	2	4	0
3	2	3	0
4	3	6	0
5	4	4	0
6	9	10	1
7	6	8	1
8	9	5	1
9	8	7	1
10	10	8	1

Using the K-Nearest Neighbors (K-NN) classification algorithm with  $k = 3$ , predict the output  $y$  for a new data point  $(x_1, x_2) = (5, 8)$ . Show all mathematical steps and calculations leading to the predicted output.

## Solution

### Step 1: Calculate Euclidean Distances

The Euclidean distance between two points  $(x_1, x_2)$  and  $(x'_1, x'_2)$  is given by:

$$d = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2}$$

We compute the Euclidean distance between the point  $(5, 8)$  and all other points in the dataset.

S. No.	$(x_1, x_2)$	Distance from (5, 8)
1	(4, 1)	$\sqrt{(5-4)^2 + (8-1)^2} = \sqrt{1+49} = \sqrt{50} \approx 7.07$
2	(2, 4)	$\sqrt{(5-2)^2 + (8-4)^2} = \sqrt{9+16} = \sqrt{25} = 5.00$
3	(2, 3)	$\sqrt{(5-2)^2 + (8-3)^2} = \sqrt{9+25} = \sqrt{34} \approx 5.83$
4	(3, 6)	$\sqrt{(5-3)^2 + (8-6)^2} = \sqrt{4+4} = \sqrt{8} \approx 2.83$
5	(4, 4)	$\sqrt{(5-4)^2 + (8-4)^2} = \sqrt{1+16} = \sqrt{17} \approx 4.12$
6	(9, 10)	$\sqrt{(5-9)^2 + (8-10)^2} = \sqrt{16+4} = \sqrt{20} \approx 4.47$
7	(6, 8)	$\sqrt{(5-6)^2 + (8-8)^2} = \sqrt{1+0} = \sqrt{1} = 1.00$
8	(9, 5)	$\sqrt{(5-9)^2 + (8-5)^2} = \sqrt{16+9} = \sqrt{25} = 5.00$
9	(8, 7)	$\sqrt{(5-8)^2 + (8-7)^2} = \sqrt{9+1} = \sqrt{10} \approx 3.16$
10	(10, 8)	$\sqrt{(5-10)^2 + (8-8)^2} = \sqrt{25+0} = \sqrt{25} = 5.00$

### Step 2: Select the 3 Nearest Neighbors

We sort the distances in ascending order:

S. No.	Distance	Class
7	1.00	1
4	2.83	0
9	3.16	1
5	4.12	0
6	4.47	1
2	5.00	0
8	5.00	1
10	5.00	1
3	5.83	0
1	7.07	0

The 3 nearest neighbours are:

S. No.	Distance	Class
7	1.00	1
4	2.83	0
9	3.16	1

### Step 3: Predict the Class

We use the majority voting from the 3 nearest neighbors. The points 7 and 9 belong to Class 1 and the point 4 belongs to Class 0, as shown in the above table. Here, Class 1 contains the highest number of points (2 points when compared with Class 0 which contains only one point). Thus, the predicted class for the point (5, 8) is Class 1.